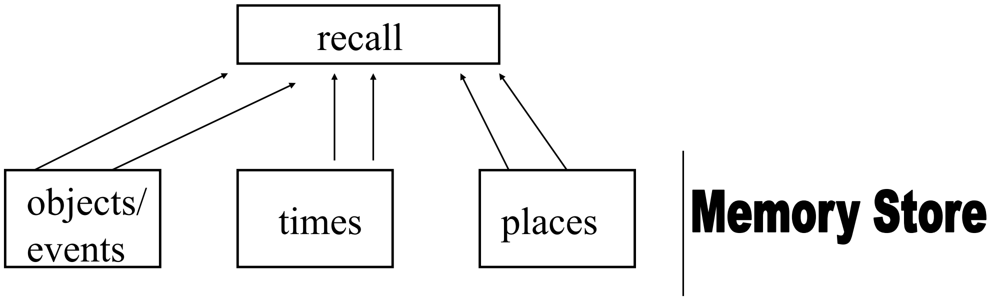
Consider this model of eyewitness memory:


recall

objects/ events

times

places

Memory Store

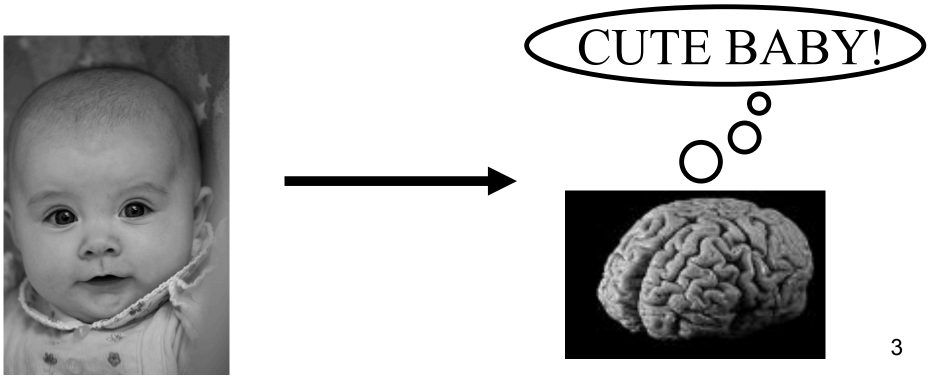
Which of the memory sins that we have discussed can this model explain (and how)? How might the model be altered to explain other sins?

Object recognition:

When perception meets memory

What is object recognition?

- Mapping between
  - Perceptual input
  - Information stored in memory



Properties of object recognition

Object recognition is:

- fast

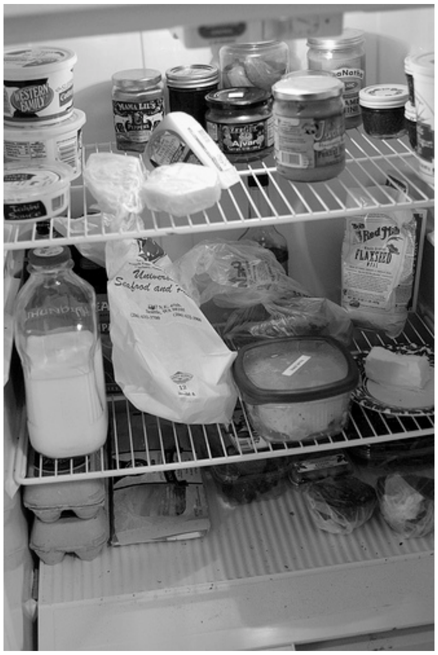
Rapid identification

Properties of object recognition

Object recognition is:

- fast
- robust

Recognition of occluded objects

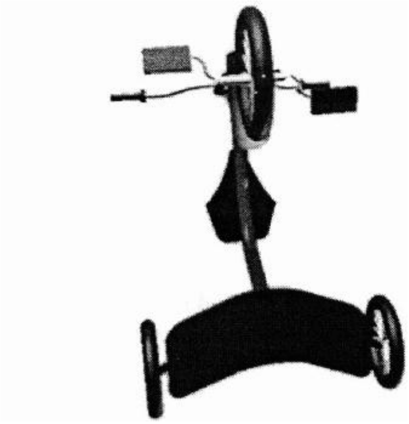


Properties of object recognition

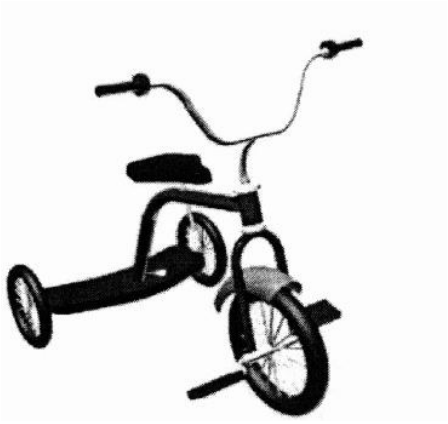
Object recognition is:

- fast
- robust
- flexible

Non-canonical view



Canonical view



Object Recognition Models

Models or theories of object recognition need to specify:

(1) the type of representation in memory
(2) the processes used to compare the perceptual input with memory

---

Object Recognition Models

Several kinds have been proposed:

- template models
- feature models
- structural models

Template models

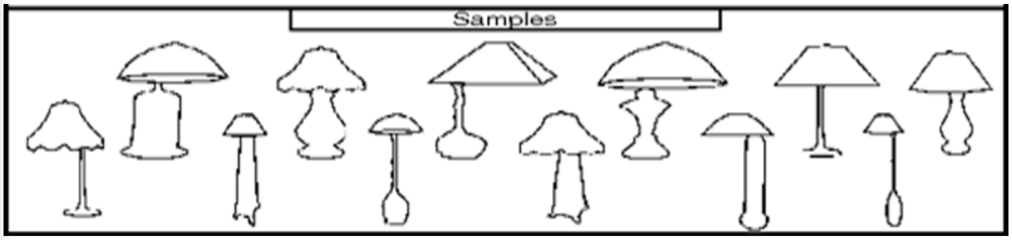
A *template* is a previously experienced visual pattern that is stored in memory.

A

Template for the letter "A"

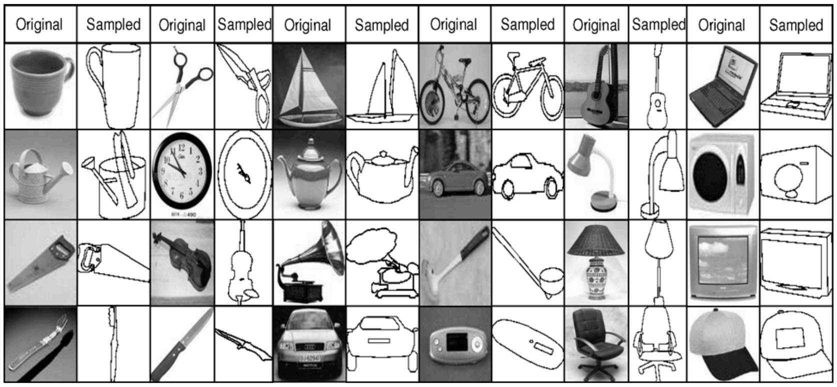
Template models

Template models assume that, through visual experience, a large number of templates are created and stored in memory.



Template matching

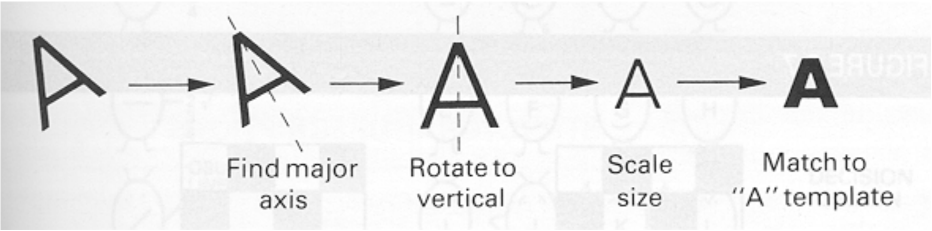
Incoming visual patterns are compared with the set of existing templates. If one matches the visual pattern, then that pattern is “recognized”.



Template matching

Because the size and orientation of objects will vary, many models assume that the image can be *transformed* during the matching process.

- – rotated
- – scaled



Feature models

Feature models assume that objects are composed of separable parts, called *features*. Object memory consists of a list of features.

Left-slanted line
Right-slanted line
Horizontal line

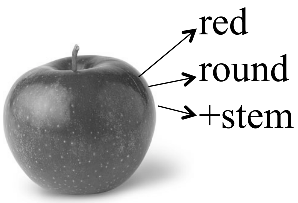
Features for the letter “A”

Feature matching

Feature models assume that during visual processing, features are extracted from the image.

Feature matching

This feature set is then compared against the feature lists stored in memory. If most features match, then the object is recognized.



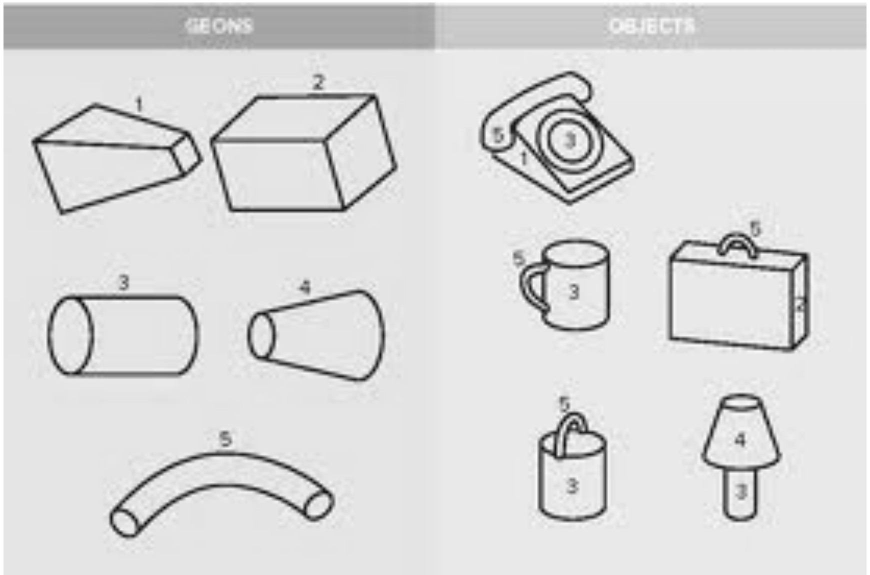
	apple	**broccoli**	**heart**	**dog**
red	x		x	
round	x			
+stem	x	x		
fluffy				x
edible	x	x		

Structural models

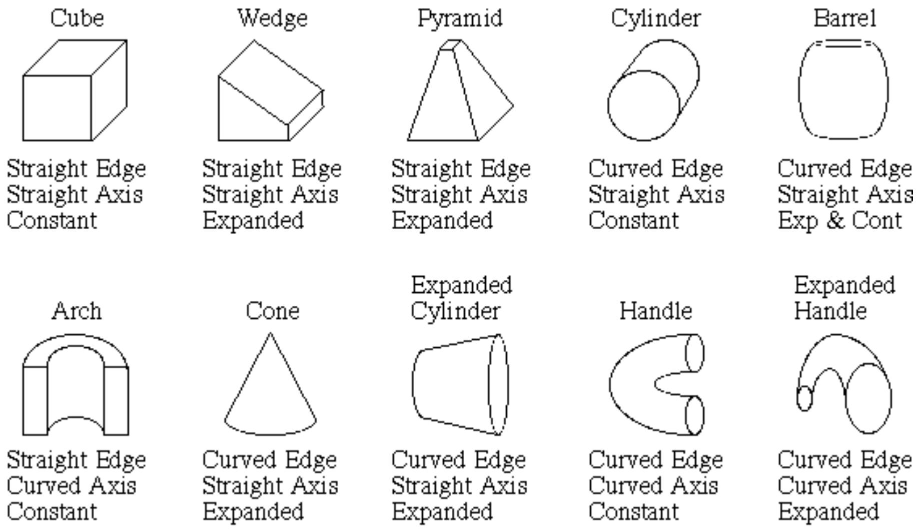
Structural models are a type of feature model that assume that *shapes and their spatial relations* are most important.

One structural model is Irving Biederman's "Recognition by Components" model. It proposes a set of basic shapes ("geons") that are used to represent more complex objects.

Geons → Objects



Repertoire of Geons



Matching in structural models

Grouping laws can be used to extract information about geons from images. These are compared with the structural descriptions in memory.

Note that structural descriptions depict spatial relationships *relative to the object* and not relative to the viewer. Thus transformations (like rotations) are not needed.

Viewer- vs. Object-centered

Two ways of representing spatial relations:

- In viewer-centered models, features are described by reference to how they appear in a specific image.

> For example, “horizontal line” in the description of the letter “A” would becomes a *vertical* line if the same image were rotated 90 degrees.

Thus, this kind of description critically depends on how an object is viewed, and multiple descriptions are needed for objects seen from different angles.

Viewer- vs. Object-centered

Two ways of representing spatial relations:

- In object-centered models, features are described in relation to other features on the same object.

> An object-centered description of the letter “A” for example might be: “two straight lines (*a, b*) connected at one end at an angle of less than 90 degrees” plus “a line (*c*) connecting the two midpoints of *a* and *b.*”

Thus, this kind of description is true no matter how the object is viewed.

Testing the models

Think of an experiment to decide between:

- GROUP 1: a template from a feature model
- GROUP 2: a viewer-centered from an object-centered (e.g., geon) feature model

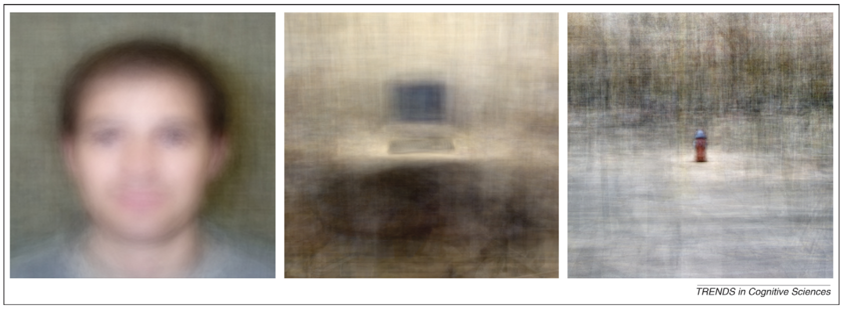
Template models

Strengths:

- Memory representation is built directly from perceptual information.
- Works well for limited numbers of fairly distinct items (e.g., numbers on checks).

Template models: Strengths

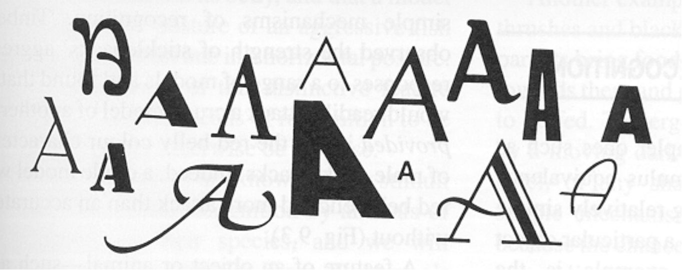
- Averaging across similar objects results in recognizable categories

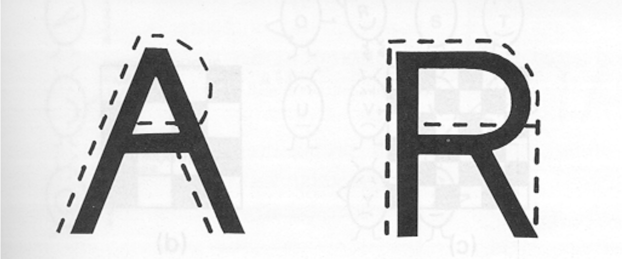


TRENDS in Cognitive Sciences

Figure 1. The structure of objects and their backgrounds. In this illustration, each image has been created by averaging hundreds of pictures containing a particular object in the center (a face, keyboard and fire hydrant) at a fixed scale and pose. Images come from the LabelMe dataset [65]. Before averaging, each picture is translated and scaled so that the target object is in the center. No other transformations are applied. The averages reveal the regularities existing in the intensity patterns across all the images. The background of many objects does not average to a uniform field, showing that an object extends its influence beyond its own boundaries, and this property is heavily used by the visual system.

Problems with template models





Template models

Weaknesses:

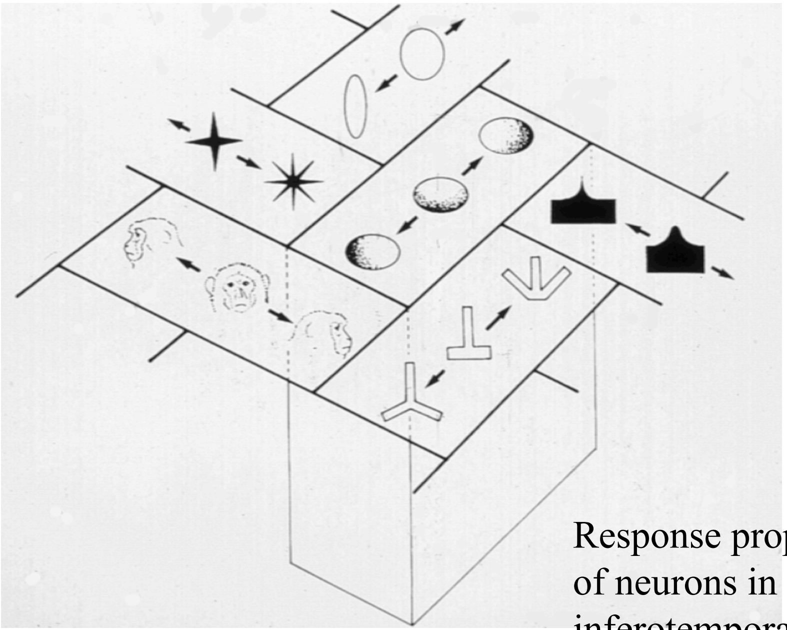
- Need *many* templates even for one object.
- Search through all the necessary templates would seem to take too much time.

Feature models (viewer-centered)

Strengths:

- Allow more flexibility than template models.
- Visual search experiments show that objects with distinct features can be found more quickly.
- Some neurobiological support for feature detection.

Feature detection in the visual system



Response properties of neurons in the inferotemporal cortex

Feature models (viewer-centered)

Weaknesses:

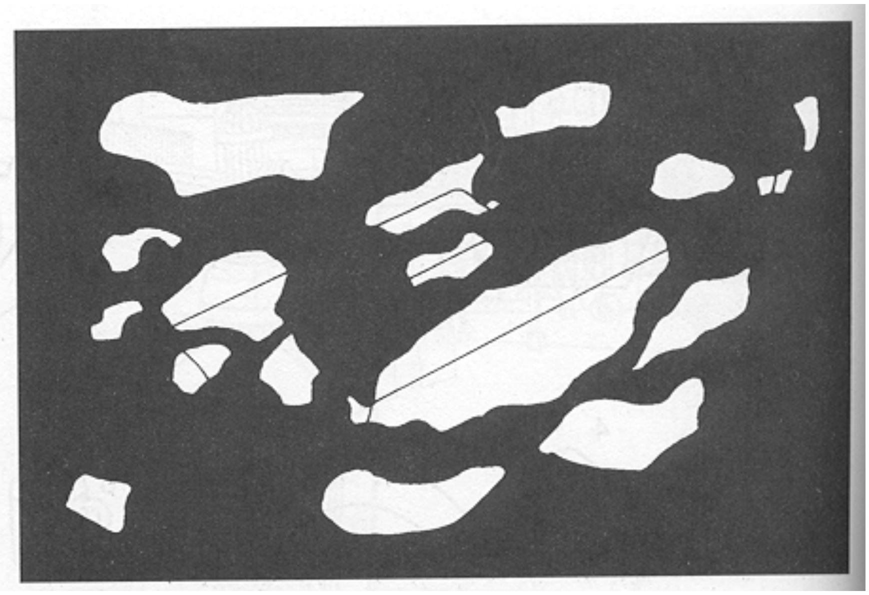
- How are the features determined?
- Still need many feature lists, for multiple views of the same object (e.g., letter A upright and upside-down)

Geon model

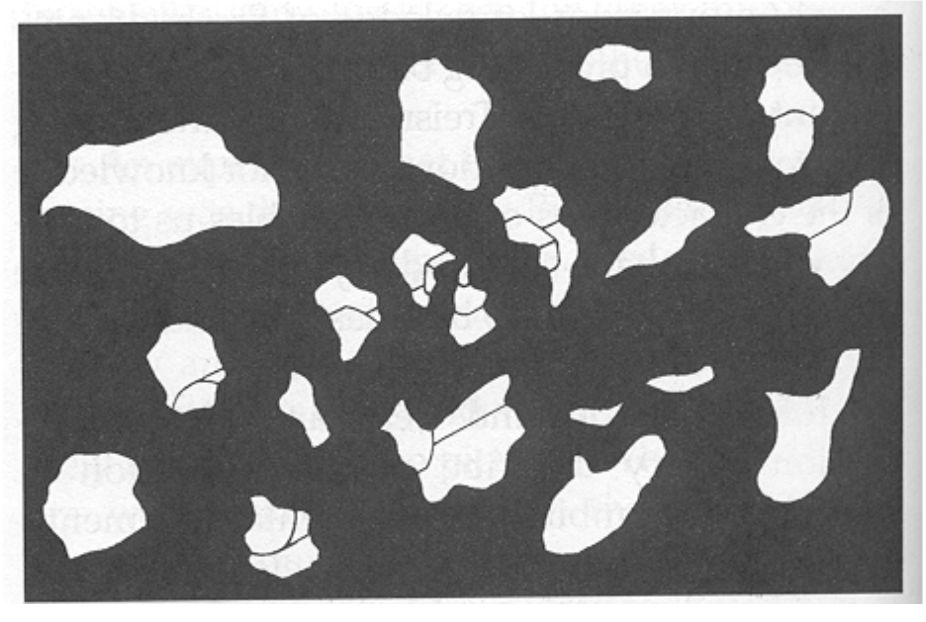
Strengths:

- Relationships between features are built in.
- We might be able to pick out features based on general grouping principles.
- Experimental data suggest that people are better at recognizing degraded objects when more of the geon information is intact.

Object with geon intersections occluded



Object with geon intersections revealed



Another example

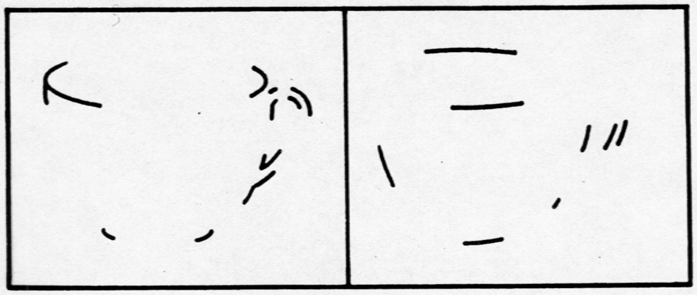


Figure 2.8 *Illustration of 65% contour removal centered at either midsegments (left object) or vertices (right object).*

Geon model

Weaknesses:

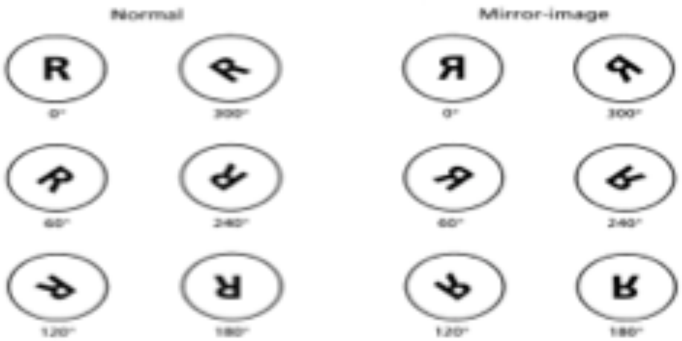
- Not all objects are well-described by geons (bread).
- Shape can’t always distinguish between two objects (orange vs. apple vs. potato)
- People show some view-point dependence.

Experimental observations

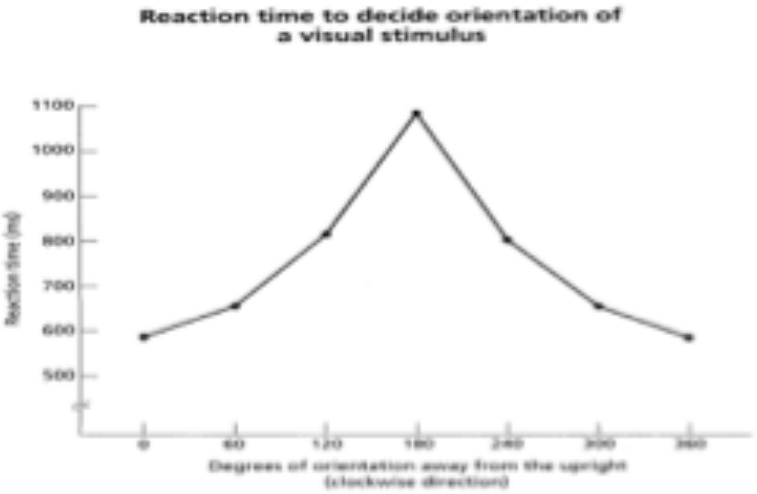
R Vs. **Я**

Experimental observations

Different degrees of rotations performed on the materials (Cooper & Shepard, 1973)

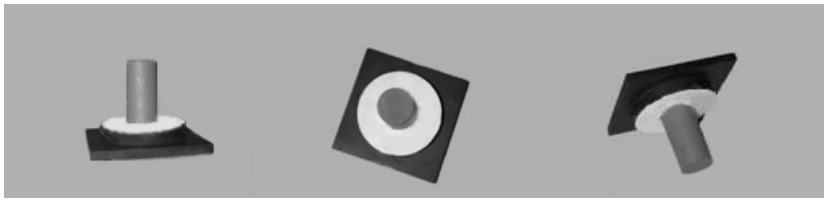


Experimental observations



Experimental observations

- Tarr and Pinker (1989)
- Train on new objects at three orientations.



  - RT is a function of amount of rotation from closest taught orientation
  - suggests that people have learned three different views of the object.

Summary

Still an active area of investigation!

- Maybe something in between:
  - descriptions in terms of features, with shape playing an important role
  - multiple descriptions of each object (from a few different viewpoints)
  - processes for transforming objects to find closest viewpoint for matching
- Maybe multiple types operate in parallel, depending on type of stimulus and level of specificity at which we need to identify it.

Thought Question

Imagine that you discover a group of patients who seem to have selectively lost the ability to recognize objects presented in “unusual” views. Their identification of objects presented in more typical views appears unchanged. Which models of object recognition is this observation consistent with? With which models is this inconsistent?